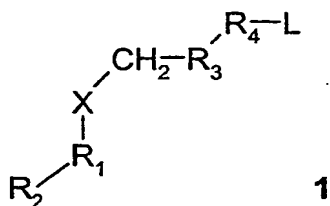


Claims

1. A compound of formula 1



wherein R<sub>1</sub>-R<sub>2</sub> is a group recognized by AGT as a substrate;

X is oxygen or sulfur;

R<sub>3</sub> is an aromatic or a heteroaromatic group, or an optionally substituted unsaturated alkyl, cycloalkyl or heterocyclyl group with the double bond connected to CH<sub>2</sub>;

R<sub>4</sub> is a linker; and

L is a label, a plurality of same or different labels, a bond connecting R<sub>4</sub> to R<sub>1</sub> forming a cyclic substrate, or a further group -R<sub>3</sub>-CH<sub>2</sub>-X-R<sub>1</sub>-R<sub>2</sub>.

2. The compound of formula 1 according to claim 1, wherein

R<sub>1</sub> is a heteroaromatic group containing 1 to 5 nitrogen atoms;

R<sub>2</sub> is hydrogen, alkyl of 1 to 10 carbon atoms, or a saccharide moiety;

X is oxygen;

R<sub>3</sub> is phenyl, an unsubstituted or substituted mono- or bicyclic heteroaryl group of 5 or 6 rings atoms comprising zero, one, two, three or four ring nitrogen atoms and zero or one oxygen atom and zero or one sulfur atom, with the proviso that at least one ring carbon atom is replaced by a nitrogen, oxygen or sulfur atom, 1-alkenyl, 1-alkinyl, 1-cyclohexenyl with 3 to 7 carbon atoms, or an optionally substituted unsaturated heterocyclyl group with 3 to 12 atoms and 1 to 5 heteroatoms selected from nitrogen, oxygen and sulfur, and a double bond in the position connecting the heterocyclyl group to methylene CH<sub>2</sub>;

R<sub>4</sub> is an optionally substituted straight or branched chain alkylene group with 1 to 300 carbon atoms, wherein optionally

(a) one or more carbon atoms are replaced by oxygen

(b) one or more carbon atoms are replaced by nitrogen carrying a hydrogen atom, and the adjacent carbon atoms are substituted by oxo, representing an amide function -NH-CO-;

(c) one or more carbon atoms are replaced by oxygen, and the adjacent carbon atoms are substituted by oxo, representing an ester function -O-CO-;

- 61 -

- (d) the bond between two adjacent carbon atoms is a double or a triple bond, representing a function  $-\text{CH}=\text{CH}-$  or  $-\text{C}\equiv\text{C}-$ ;
- (e) one or more carbon atoms are replaced by a phenylene, a saturated or unsaturated cycloalkylene, a saturated or unsaturated bicycloalkylene, a bridging heteroaromatic or a
- 5 bridging saturated or unsaturated heterocyclyl group; and/or
- (f) two adjacent carbon atoms are replaced by a disulfide linkage  $-\text{S}-\text{S}-$ ; and

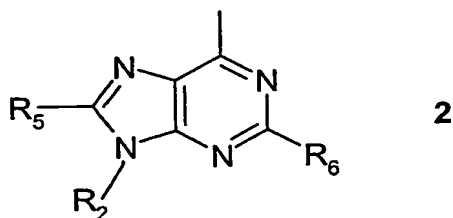
L is one or a plurality of same or different labels selected from a spectroscopic probe, a magnetic probe, a contrast reagent, a molecule which is one part of a specific binding pair

10 which is capable of specifically binding to a partner, a molecule that is suspected to interact with other biomolecules, a library of molecules that are suspected to interact with other biomolecules, a molecule which is capable of crosslinking to other molecules, a molecule which is capable of generating hydroxyl radicals upon exposure to  $\text{H}_2\text{O}_2$  and ascorbate, a molecule which is capable of generating reactive radicals upon irradiation with light, a

15 molecule covalently attached to a solid support, a nucleic acid or a derivative thereof capable of undergoing base-pairing with its complementary strand, a lipid or other hydrophobic molecule with membrane-inserting properties, a biomolecule with desirable enzymatic, chemical or physical properties, a bond connecting  $\text{R}_4$  to  $\text{R}_1$  forming a cyclic substrate, and a further group  $-\text{R}_3-\text{CH}_2-\text{X}-\text{R}_1-\text{R}_2$ .

20

3. The compound of formula 1 according to claim 1, wherein  $\text{R}_1-\text{R}_2$  is a radical of formula 2



wherein  $\text{R}_2$  is hydrogen, alkyl of 1 to 10 carbon atoms, or a saccharide moiety;

$\text{R}_5$  is hydrogen, halogen, trifluoromethyl, or hydroxy; and

25  $\text{R}_6$  is hydrogen, hydroxy or unsubstituted or substituted amino; and tautomeric forms thereof.

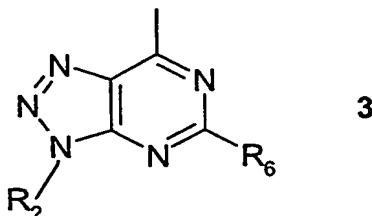
4. The compound of formula 1 according to claim 3, wherein the saccharide moiety  $\text{R}_2$  is a  $\beta$ -D-2'-deoxyribosyl, or a  $\beta$ -D-2'-deoxyribosyl being incorporated into a single stranded

30 oligodeoxyribonucleotide having a length of 2 to 99 nucleotides, wherein the guanine derivative  $\text{R}_1$  occupies any position within the oligonucleotide sequence.

- 62 -

5. The compound of formula 1 according to claim 3 wherein  $R_2$  is hydrogen,  $R_5$  is hydrogen,  $R_6$  is unsubstituted amino, and X is oxygen.

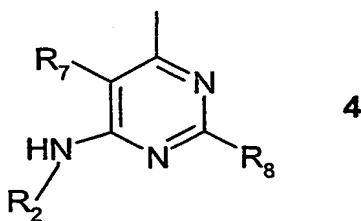
6. The compound of formula 1 according to claim 1, wherein  $R_1$ - $R_2$  is a radical of formula 3



5

wherein  $R_2$  is hydrogen, alkyl of 1 to 10 carbon atoms, or a saccharide moiety; and  $R_6$  is hydrogen, hydroxy or unsubstituted or substituted amino; and tautomeric forms thereof.

10 7. The compound of formula 1 according to claim 1, wherein  $R_1$ - $R_2$  is a radical of formula 4



wherein  $R_2$  is hydrogen, alkyl of 1 to 10 carbon atoms, or a saccharide moiety; and  $R_7$  and  $R_8$  are both independently of one another hydrogen, halogen, lower alkyl with 1 to 4 carbon atoms, amino, or nitro.

15

8. The compound of formula 1 according to claim 1 wherein  $R_3$  is triazolyl, tetrazolyl, isoxazolyl, thienyl, or isoxazolidinyl.

9. The compound of formula 1 according to claim 8 wherein  $R_3$  is triazolyl.

20

10. The compound of formula 1 according to claim 8 wherein  $R_3$  is tetrazolyl.

11. The compound of formula 1 according to claim 8 wherein  $R_3$  is isoxazolyl.

25

12. The compound of formula 1 according to claim 8 wherein  $R_3$  is thienyl.

13. The compound of formula 1 according to claim 8 wherein  $R_3$  is isoxazolidinyl.

14. The compound of formula 1 according to claim 1 wherein  $R_3$  is 1-alkynyl.

15. The compound of formula 1 according to claim 1 wherein  $R_4$  is a straight chain alkylene group with 2 to 25 carbon atoms, a straight chain polyethylene glycol group with 4 to 100 ethyleneoxy units, or a straight chain alkylene group with 2 to 25 carbon atoms wherein two or more carbon atoms are replaced by an amide function  $-NH-CO-$ , optionally attached to the group  $R_3$  by a  $-CH=CH-$  or  $-C\equiv C-$  group.

16. The compound of formula 1 according to claim 1 wherein  $R_4$  is a branched chain alkylene group comprising a polyethylene glycol group of 3 to 6 ethylene glycol units and alkylene groups wherein carbon atoms are replaced by amide bonds, and further carrying substituted amino and hydroxy functions.

17. The compound of formula 1 according to claim 1 wherein  $R_4$  is a branched chain alkylene group having a dendritic structures wherein amine, carboxamide and ether functions replace carbon atoms of the alkylene group.

18. The compound of formula 1 according to claim 3 wherein  $R_3$  is phenylene and L is a further group  $-R_3-CH_2-X-R_1-R_2$ .

19. The compound of formula 1 according to claim 3 or 18 wherein  $R_3$  is 1,4-phenylene and linker  $R_4$  is a straight chain alkylene group of 10 to 40 carbon atoms wherein 3 to 12 carbon atoms are replaced by oxygen, one or two carbon atoms are replaced by 1,4-triazolidene units, and optionally one carbon atom is replaced by a 1,4-phenylene unit.

20. The compound of formula 1 according to claim 3 or 18 wherein  $R_3$  is 1,4-phenylene and linker  $R_4$  is a straight chain alkylene group of 10 to 40 carbon atoms optionally substituted by oxo wherein 3 to 12 carbon atoms are replaced by oxygen and one or two carbon atoms are replaced by nitrogen.

21. The compound of formula 1 according to claim 3 or 18 wherein  $R_3$  is 1,4-phenylene and linker  $R_4$  is a straight chain alkylene group of 6 to 40 carbon atoms wherein 2 to 12 carbon atoms are replaced by oxygen and one or two bonds between two adjacent carbon atoms is a double bond.

22. The compound of formula 1 according to claim 3 wherein  $R_3$  is phenylene,  $R_6$  is amino and L is a bond connecting  $R_4$  to  $R_6$ .

23. The compound of formula 1 according to claim 3 wherein  $R_3$  is phenylene and L is methotrexate.

5 24. The compound of formula 1 according to claim 3 wherein  $R_3$  is phenylene and L is a plurality of same or different labels.

25. The compound of formula 1 according to claim 24 wherein  $R_3$  is phenylene and L is two different labels.

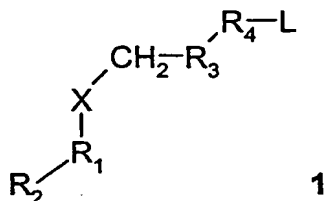
10

26. A method for the synthesis of a compound of the formula 1 according to any one of claims 1 to 25.

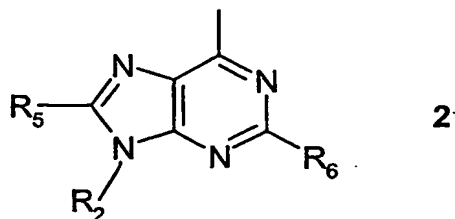
15 27. A method according to claim 26, characterized in that a compound of the formula  $R_2-R_1-X-CH_2-R_3-R_4$  wherein  $R_4$  is a polyfunctional residue having two or more reactive nucleophilic or electrophilic groups is reacted with a suitable reagent introducing one or more labels L.

20 28. A method according to claim 27 wherein the reactive nucleophilic or electrophilic groups are orthogonally protected functional groups, and conditions are chosen such as to allow separate deprotection so that each released functionality in turn can be further chemically manipulated either to attach a label to it or for the introduction of further extension of linker  $R_4$ .

25 29. A compound of the formula 1



wherein  $R_1-R_2$  is a radical of formula 2



- 65 -

wherein  $R_2$  is hydrogen,  $R_5$  is hydrogen and  $R_6$  is unsubstituted amino;

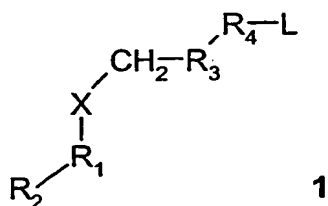
X is oxygen;

$R_3$  is triazolyl, tetrazolyl, isoxazolyl, thienyl, isoxazolidinyl or alkynyl;

$R_4$  is a linker; and

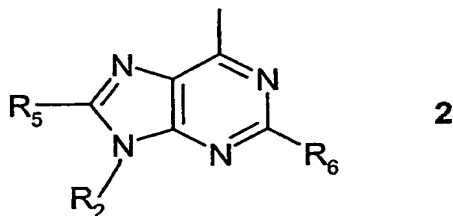
5 L is amino or azido.

30. A compound of the formula 1



10

wherein  $R_1$ - $R_2$  is a radical of formula 2



wherein  $R_2$  is hydrogen,  $R_5$  is hydrogen and  $R_6$  is unsubstituted amino;

X is oxygen;

15  $R_3$  is 1,4-phenylene;

$R_4$  is a straight chain alkylene group of 10 to 40 carbon atoms optionally substituted by oxo wherein up to 12 carbon atoms are replaced by oxygen and zero, one or two carbon atoms are replaced by nitrogen; and

L is amino or azido.

20

31. A method for detecting and manipulating a protein of interest, characterized in that the protein of interest incorporated into a AGT fusion protein is contacted with an AGT substrates carrying a label, and the AGT fusion protein is detected and optionally further manipulated using the label in a system designed for recognising or handling the label, and wherein the

25 AGT substrate carrying the label is a compound of formula 1 according to any one of claims 1 to 25.